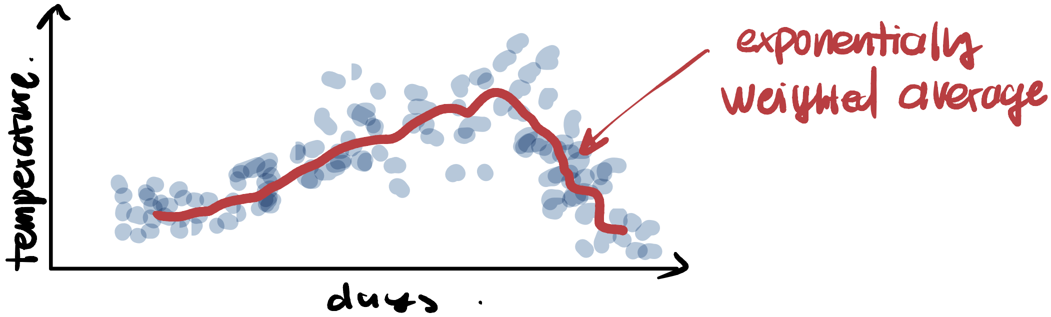
# Exponentially weighted averages

Example of exponentially weighted averages. (Temperature in London).

$\theta_1 = 40°F$
$\theta_2 = 49°F$
$\theta_3 = 45°F$
$\vdots$
$\theta_{356} = 56°F$.




$$V_t = \underline{\beta} V_{t-1} + (1-\beta)\theta_t$$

$\beta$ is a hyperparameter:

$\beta = 0.9$ : $\approx$ 10 days temperature avg.

$\beta = 0.98$ : $\approx$ 50 days temperature avg.

$\beta = 0.2$ : $\approx$ 2 days

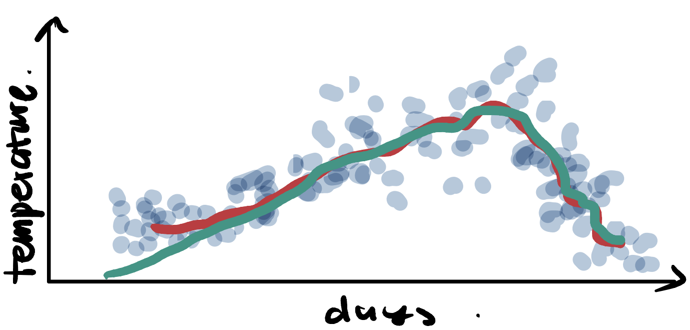
$V_t$ is approximately average over $\rightarrow \approx \frac{1}{1-\beta}$ days

e.g. $\frac{1}{1-0.98} \approx \underline{50}$

Decreasing $\beta$ will create more oscilliation (more sensitive to data changes).

Increasing $\beta$ will shift $V_t$ slightly to right (longer latency).

## Bias corrction.




The green line is what we actually get by applying $V_t = \beta V_{t-1} + (1-\beta)\theta_t$ for each $\theta_t$. Green line starts lower than the actuall exponentially weighted average and diverge with red line later on.

$V_0 = 0$

$V_1 = 0.98 V_0 + 0.02\theta_1 = \underline{0.02\theta_1}$

$V_2 = 0.98 V_1 + 0.02\theta_2 = \underline{0.0196\theta_1 + 0.02\theta_2}$

Starts with very small values.

Bias correction helps to adjust the slow startup phase.

$\longrightarrow \quad V_t = \dfrac{V_t}{1-\beta^t}$ , e.g. $t=2$: $1-\beta^t = 1-(0.98)^2 = 0.0396$

$$\frac{V_2}{0.0396} = \frac{0.0196\theta_1 + 0.02\theta_2}{0.0396}$$

scale up $\theta_1$ and $\theta_2$

Since $\beta < 0$, $\lim_{t\to\infty} \beta^t = 0$ , $V_t = \dfrac{V_t}{1-0} = V_t$.

Thus the bias correction will not affect $V_t$

after the startup phase.